

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/659,055A
Source: IFW/6
Date Processed by STIC: 6/21/06

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:02

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

3 <110> APPLICANT: Takeda San Diego, Inc.
 5 <120> TITLE OF INVENTION: CRYSTALLIZATION OF DIPEPTIDYL PEPTIDASE IV (DPPIV)
 7 <130> FILE REFERENCE: DPPIV-5001-C1
 9 <140> CURRENT APPLICATION NUMBER: 10/659,055A
 10 <141> CURRENT FILING DATE: 2003-09-09
 12 <150> PRIOR APPLICATION NUMBER: US 60/409,206
 13 <151> PRIOR FILING DATE: 2002-09-09
 15 <160> NUMBER OF SEQ ID NOS: 3
 17 <170> SOFTWARE: PatentIn version 3.3
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 766
 21 <212> TYPE: PRT
 22 <213> ORGANISM: Homo sapiens
 25 <220> FEATURE:
 W--> 26 <221> NAME/KEY: Amino acid sequence for full-length human wild type DPPIV
 27 <222> LOCATION: (1)..(766)
 29 <300> PUBLICATION INFORMATION:
 30 <308> DATABASE ACCESSION NO: Genbank/NP_001926
 31 <309> DATABASE ENTRY DATE: 2002-02-19
 32 <313> RELEVANT RESIDUES: (1)..(766)
 34 <400> SEQUENCE: 1
 36 Met Lys Thr Pro Trp Lys Val Leu Leu Gly Leu Leu Gly Ala Ala Ala
 37 1 5 10 15
 40 Leu Val Thr Ile Ile Thr Val Pro Val Val Leu Leu Asn Lys Gly Thr
 41 20 25 30
 44 Asp Asp Ala Thr Ala Asp Ser Arg Lys Thr Tyr Thr Leu Thr Asp Tyr
 45 35 40 45
 48 Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr Ser Leu Arg Trp Ile Ser
 49 50 55 60
 52 Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn Asn Ile Leu Val Phe Asn
 53 65 70 75 80
 56 Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu Glu Asn Ser Thr Phe Asp
 57 85 90 95
 60 Glu Phe Gly His Ser Ile Asn Asp Tyr Ser Ile Ser Pro Asp Gly Gln
 61 100 105 110
 64 Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys Gln Trp Arg His Ser Tyr
 65 115 120 125
 68 Thr Ala Ser Tyr Asp Ile Tyr Asp Leu Asn Lys Arg Gln Leu Ile Thr
 69 130 135 140
 72 Glu Glu Arg Ile Pro Asn Asn Thr Gln Trp Val Thr Trp Ser Pro Val
 73 145 150 155 160
 76 Gly His Lys Leu Ala Tyr Val Trp Asn Asn Asp Ile Tyr Val Lys Ile
 77 165 170 175

RAW SEQUENCE LISTING

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:02

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

```

80 Glu Pro Asn Leu Pro Ser Tyr Arg Ile Thr Trp Thr Gly Lys Glu Asp
81      180      185      190
84 Ile Ile Tyr Asn Gly Ile Thr Asp Trp Val Tyr Glu Glu Glu Val Phe
85      195      200      205
88 Ser Ala Tyr Ser Ala Leu Trp Trp Ser Pro Asn Gly Thr Phe Leu Ala
89      210      215      220
92 Tyr Ala Gln Phe Asn Asp Thr Glu Val Pro Leu Ile Glu Tyr Ser Phe
93 225      230      235      240
96 Tyr Ser Asp Glu Ser Leu Gln Tyr Pro Lys Thr Val Arg Val Pro Tyr
97      245      250      255
100 Pro Lys Ala Gly Ala Val Asn Pro Thr Val Lys Phe Phe Val Val Asn
101      260      265      270
104 Thr Asp Ser Leu Ser Ser Val Thr Asn Ala Thr Ser Ile Gln Ile Thr
105      275      280      285
108 Ala Pro Ala Ser Met Leu Ile Gly Asp His Tyr Leu Cys Asp Val Thr
109      290      295      300
112 Trp Ala Thr Gln Glu Arg Ile Ser Leu Gln Trp Leu Arg Arg Ile Gln
113 305      310      315      320
116 Asn Tyr Ser Val Met Asp Ile Cys Asp Tyr Asp Glu Ser Ser Gly Arg
117      325      330      335
120 Trp Asn Cys Leu Val Ala Arg Gln His Ile Glu Met Ser Thr Thr Gly
121      340      345      350
124 Trp Val Gly Arg Phe Arg Pro Ser Glu Pro His Phe Thr Leu Asp Gly
125      355      360      365
128 Asn Ser Phe Tyr Lys Ile Ile Ser Asn Glu Glu Gly Tyr Arg His Ile
129      370      375      380
132 Cys Tyr Phe Gln Ile Asp Lys Lys Asp Cys Thr Phe Ile Thr Lys Gly
133 385      390      395      400
136 Thr Trp Glu Val Ile Gly Ile Glu Ala Leu Thr Ser Asp Tyr Leu Tyr
137      405      410      415
140 Tyr Ile Ser Asn Glu Tyr Lys Gly Met Pro Gly Gly Arg Asn Leu Tyr
141      420      425      430
144 Lys Ile Gln Leu Ser Asp Tyr Thr Lys Val Thr Cys Leu Ser Cys Glu
145      435      440      445
148 Leu Asn Pro Glu Arg Cys Gln Tyr Tyr Ser Val Ser Phe Ser Lys Glu
149      450      455      460
152 Ala Lys Tyr Tyr Gln Leu Arg Cys Ser Gly Pro Gly Leu Pro Leu Tyr
153 465      470      475      480
156 Thr Leu His Ser Ser Val Asn Asp Lys Gly Leu Arg Val Leu Glu Asp
157      485      490      495
160 Asn Ser Ala Leu Asp Lys Met Leu Gln Asn Val Gln Met Pro Ser Lys
161      500      505      510
164 Lys Leu Asp Phe Ile Ile Leu Asn Glu Thr Lys Phe Trp Tyr Gln Met
165      515      520      525
168 Ile Leu Pro Pro His Phe Asp Lys Ser Lys Lys Tyr Pro Leu Leu Leu
169      530      535      540
172 Asp Val Tyr Ala Gly Pro Cys Ser Gln Lys Ala Asp Thr Val Phe Arg
173 545      550      555      560
176 Leu Asn Trp Ala Thr Tyr Leu Ala Ser Thr Glu Asn Ile Ile Val Ala

```

RAW SEQUENCE LISTING

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:02

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

```

177          565          570          575
180 Ser Phe Asp Gly Arg Gly Ser Gly Tyr Gln Gly Asp Lys Ile Met His
181          580          585          590
184 Ala Ile Asn Arg Arg Leu Gly Thr Phe Glu Val Glu Asp Gln Ile Glu
185          595          600          605
188 Ala Ala Arg Gln Phe Ser Lys Met Gly Phe Val Asp Asn Lys Arg Ile
189          610          615          620
192 Ala Ile Trp Gly Trp Ser Tyr Gly Gly Tyr Val Thr Ser Met Val Leu
193 625          630          635          640
196 Gly Ser Gly Ser Gly Val Phe Lys Cys Gly Ile Ala Val Ala Pro Val
197          645          650          655
200 Ser Arg Trp Glu Tyr Tyr Asp Ser Val Tyr Thr Glu Arg Tyr Met Gly
201          660          665          670
204 Leu Pro Thr Pro Glu Asp Asn Leu Asp His Tyr Arg Asn Ser Thr Val
205          675          680          685
208 Met Ser Arg Ala Glu Asn Phe Lys Gln Val Glu Tyr Leu Leu Ile His
209          690          695          700
212 Gly Thr Ala Asp Asp Asn Val His Phe Gln Gln Ser Ala Gln Ile Ser
213 705          710          715          720
216 Lys Ala Leu Val Asp Val Gly Val Asp Phe Gln Ala Met Trp Tyr Thr
217          725          730          735
220 Asp Glu Asp His Gly Ile Ala Ser Ser Thr Ala His Gln His Ile Tyr
221          740          745          750
224 Thr His Met Ser His Phe Ile Lys Gln Cys Phe Ser Leu Pro
225          755          760          765

```

228 <210> SEQ ID NO: 2

229 <211> LENGTH: 2184

230 <212> TYPE: DNA

231 <213> ORGANISM: Homo sapiens

234 <220> FEATURE:

W--> 235 <221> NAME/KEY: Human cDNA sequence encoding residues 39-766 of DPPIV

236 <222> LOCATION: (1)..(2184)

238 <400> SEQUENCE: 2

```

239 agtcgcaaaa cttacactct aactgattac ttaaaaaata cttatagact gaagttatac      60
241 tccttaagat ggatttcaga tcatgaatat ctctacaaac aagaaaataa tatcttggtg      120
243 ttcaatgctg aatatggaaa cagctcagtt ttcttgagaga acagtacatt tgatgagttt      180
245 ggacattcta tcaatgatta ttcaatatct cctgatgggc agtttattct cttagaatac      240
247 aactacgtga agcaatggag gcattcctac acagcttcat atgacattta tgatttaaata      300
249 aaaaggcagc tgattacaga agagaggatt ccaaacaaca cacagtgggt cacatggtca      360
251 ccagtgggtc ataaattggc atatgtttgg aacaatgaca tttatgttaa aattgaacca      420
253 aatttaccaa gttacagaat cacatggacg gggaaagaag atataatata taatggaata      480
255 actgactggg tttatgaaga ggaagtcttc agtgcctact ctgctctgtg gtgggtctcca      540
257 aacggcactt ttttagcata tgcccaattt aacgacacag aagtccactt tattgaatac      600
259 tccttctact ctgatgagtc actgcagtac ccaaagactg tacgggttcc atatccaaag      660
261 gcaggagctg tgaatccaac tgtaaagttc tttgttgtaa atacagactc tctcagctca      720
263 gtcaccaatg caacttccat acaaatcact gctcctgctt ctatgttgat aggggatcac      780
265 tacttggtgt atgtgacatg ggcaacacaa gaaagaattt ctttgcagtg gctcaggagg      840
267 attcagaact attcgggtcat ggatatttgt gactatgatg aatccagtgg aagatggaac      900
269 tgcttagtgtg cacggcaaca cattgaaatg agtactactg gctgggttgg aagatttagg      960

```

RAW SEQUENCE LISTING

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:02

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

```

271 ccttcagaac ctcattttac ccttgatggt aatagcttct acaagatcat cagcaatgaa 1020
273 gaagggttaca gacacatttg ctattttccaa atagataaaa aagactgcac atttattaca 1080
275 aaaggcacct gggaagtcac cgggatagaa gctctaacca gtgattatct atactacatt 1140
277 agtaatgaat ataaaggaat gccaggagga aggaatcttt ataaaatcca acttattgac 1200
279 tatacaaaag tgacatgcct cagttgtgag ctgaatccgg aaagggtgtca gtactattct 1260
281 gtgtcattca gtaaagaggc gaagtattat cagctgagat gttccgggtcc tgggtctgccc 1320
283 ctctatactc tacacagcag cgtgaatgat aaagggctga gagtcctgga agacaattca 1380
285 gctttggata aaatgctgca gaatgtccag atgccctcca aaaaactgga cttcattatt 1440
287 ttgaatgaaa caaaattttg gtatcagatg atcttgacct ctcattttga taaatccaag 1500
289 aaatatcctc tactattaga tgtgtatgca ggcccatgta gtcaaaaagc agacactgtc 1560
291 ttcagactga actgggccac ttaccttgca agcacagaaa acattatagt agctagcttt 1620
293 gatggcagag gaagtggta ccaaggagat aagatcatgc atgcaatcaa cagaagactg 1680
295 ggaacatttg aagttgaaga tcaaattgaa gcagccagac aattttcaaa aatgggattt 1740
297 gtggacaaca aacgaattgc aatttggggc tgggtcatatg gaggggtacgt aacctcaatg 1800
299 gtcctgggag cggaagtgg cgtgttcaag tgtggaatag ccgtggcgcc tgtatcccg 1860
301 tgggagtact atgactcagt gtacacagaa cgttacatgg gtctcccaac tccagaagac 1920
303 aaccttgacc attacagaaa ttcaacagtc atgagcagag ctgaaaattt taaacaagtt 1980
305 gagtacctcc ttattcatgg aacagcagat gataacgttc actttcagca gtcagctcag 2040
307 atctccaaag ccctggtcga tgttggagtg gatttccagg caatgtggta tactgatgaa 2100
309 gaccatggaa tagctagcag cacagcacac caacatatat ataccacat gagccacttc 2160
311 ataaaacaat gtttctcttt acct 2184

```

314 <210> SEQ ID NO: 3

315 <211> LENGTH: 740

316 <212> TYPE: PRT

317 <213> ORGANISM: Homo sapiens

320 <220> FEATURE:

W--> 321 <221> NAME/KEY: N-terminal tag including a partial gp67 signal sequence and

W--> 322 6x-histidine

323 <222> LOCATION: (1)..(12)

325 <220> FEATURE:

W--> 326 <221> NAME/KEY: Amino acid sequence for residues 39-766 of DPPIV with an

W--> 327 N-terminal tag

328 <222> LOCATION: (13)..(740)

330 <400> SEQUENCE: 3

332 Ala Asp Pro Gly Gly Ser His His His His His His Ser Arg Lys Thr

333 1 5 10 15

336 Tyr Thr Leu Thr Asp Tyr Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr

337 20 25 30

340 Ser Leu Arg Trp Ile Ser Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn

341 35 40 45

344 Asn Ile Leu Val Phe Asn Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu

345 50 55 60

348 Glu Asn Ser Thr Phe Asp Glu Phe Gly His Ser Ile Asn Asp Tyr Ser

349 65 70 75 80

352 Ile Ser Pro Asp Gly Gln Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys

353 85 90 95

356 Gln Trp Arg His Ser Tyr Thr Ala Ser Tyr Asp Ile Tyr Asp Leu Asn

357 100 105 110

360 Lys Arg Gln Leu Ile Thr Glu Glu Arg Ile Pro Asn Asn Thr Gln Trp

RAW SEQUENCE LISTING

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:02

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

```

361          115          120          125
364 Val Thr Trp Ser Pro Val Gly His Lys Leu Ala Tyr Val Trp Asn Asn
365          130          135          140
368 Asp Ile Tyr Val Lys Ile Glu Pro Asn Leu Pro Ser Tyr Arg Ile Thr
369 145          150          155          160
372 Trp Thr Gly Lys Glu Asp Ile Ile Tyr Asn Gly Ile Thr Asp Trp Val
373          165          170          175
376 Tyr Glu Glu Glu Val Phe Ser Ala Tyr Ser Ala Leu Trp Trp Ser Pro
377          180          185          190
380 Asn Gly Thr Phe Leu Ala Tyr Ala Gln Phe Asn Asp Thr Glu Val Pro
381          195          200          205
384 Leu Ile Glu Tyr Ser Phe Tyr Ser Asp Glu Ser Leu Gln Tyr Pro Lys
385          210          215          220
388 Thr Val Arg Val Pro Tyr Pro Lys Ala Gly Ala Val Asn Pro Thr Val
389 225          230          235          240
392 Lys Phe Phe Val Val Asn Thr Asp Ser Leu Ser Ser Val Thr Asn Ala
393          245          250          255
396 Thr Ser Ile Gln Ile Thr Ala Pro Ala Ser Met Leu Ile Gly Asp His
397          260          265          270
400 Tyr Leu Cys Asp Val Thr Trp Ala Thr Gln Glu Arg Ile Ser Leu Gln
401          275          280          285
404 Trp Leu Arg Arg Ile Gln Asn Tyr Ser Val Met Asp Ile Cys Asp Tyr
405          290          295          300
408 Asp Glu Ser Ser Gly Arg Trp Asn Cys Leu Val Ala Arg Gln His Ile
409 305          310          315          320
412 Glu Met Ser Thr Thr Gly Trp Val Gly Arg Phe Arg Pro Ser Glu Pro
413          325          330          335
416 His Phe Thr Leu Asp Gly Asn Ser Phe Tyr Lys Ile Ile Ser Asn Glu
417          340          345          350
420 Glu Gly Tyr Arg His Ile Cys Tyr Phe Gln Ile Asp Lys Lys Asp Cys
421          355          360          365
424 Thr Phe Ile Thr Lys Gly Thr Trp Glu Val Ile Gly Ile Glu Ala Leu
425          370          375          380
428 Thr Ser Asp Tyr Leu Tyr Tyr Ile Ser Asn Glu Tyr Lys Gly Met Pro
429 385          390          395          400
432 Gly Gly Arg Asn Leu Tyr Lys Ile Gln Leu Ser Asp Tyr Thr Lys Val
433          405          410          415
436 Thr Cys Leu Ser Cys Glu Leu Asn Pro Glu Arg Cys Gln Tyr Tyr Ser
437          420          425          430
440 Val Ser Phe Ser Lys Glu Ala Lys Tyr Tyr Gln Leu Arg Cys Ser Gly
441          435          440          445
444 Pro Gly Leu Pro Leu Tyr Thr Leu His Ser Ser Val Asn Asp Lys Gly
445          450          455          460
448 Leu Arg Val Leu Glu Asp Asn Ser Ala Leu Asp Lys Met Leu Gln Asn
449 465          470          475          480
452 Val Gln Met Pro Ser Lys Lys Leu Asp Phe Ile Ile Leu Asn Glu Thr
453          485          490          495
456 Lys Phe Trp Tyr Gln Met Ile Leu Pro Pro His Phe Asp Lys Ser Lys
457          500          505          510

```

VERIFICATION SUMMARY

DATE: 06/21/2006

PATENT APPLICATION: US/10/659,055A

TIME: 15:16:03

Input Set : E:\DPPIV-5001-C1 PatentIn Replacement Sequence.ST25.txt

Output Set: N:\CRF4\06212006\J659055A.raw

L:26 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:235 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:2
L:321 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:322 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:3
L:326 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:327 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:3